

# Polyelectrostatic interactions of disordered ligands suggest a physical basis for ultrasensitivity

Mikael Borg<sup>\*†‡</sup>, Tanja Mittag<sup>‡</sup>, Tony Pawson<sup>†§¶</sup>, Mike Tyers<sup>†§</sup>, Julie D. Forman-Kay<sup>\*‡</sup>, and Hue Sun Chan<sup>\*†¶</sup>

Departments of <sup>\*</sup>Biochemistry and <sup>†</sup>Medical Genetics and Microbiology, Faculty of Medicine, University of Toronto, Toronto, ON, Canada M5S 1A8; <sup>‡</sup>Molecular Structure and Function, Hospital for Sick Children, Toronto, ON, Canada M5G 1X8; and <sup>§</sup>Samuel Lunenfeld Research Institute, Mount Sinai Hospital, Toronto, ON, Canada M5G 1X5

Contributed by Tony Pawson, March 21, 2007 (sent for review February 19, 2007)

Regulation of biological processes often involves phosphorylation of intrinsically disordered protein regions, thereby modulating protein interactions. Initiation of DNA replication in yeast requires elimination of the cyclin-dependent kinase inhibitor Sic1 via the SCF<sup>Cdc4</sup> ubiquitin ligase. Intriguingly, the substrate adapter subunit Cdc4 binds to Sic1 only after phosphorylation of a minimum of any six of the nine cyclin-dependent kinase sites on Sic1. To investigate the physical basis of this ultrasensitive interaction, we consider a mean-field statistical mechanical model for the electrostatic interactions between a single receptor site and a conformationally disordered polyvalent ligand. The formulation treats phosphorylation sites as negative contributions to the total charge of the ligand and addresses its interplay with the strength of the favorable ligand–receptor contact. Our model predicts a threshold number of phosphorylation sites for receptor–ligand binding, suggesting that ultrasensitivity in the Sic1–Cdc4 system may be driven at least in part by cumulative electrostatic interactions. This hypothesis is supported by experimental affinities of Cdc4 for Sic1 fragments with different total charges. Thus, polyelectrostatic interactions may provide a simple yet powerful framework for understanding the modulation of protein interactions by multiple phosphorylation sites in disordered protein regions.

cooperativity | cyclin-dependent kinase inhibitor | DNA replication | phosphorylation | yeast cell cycle

Cellular regulatory networks have evolved finely tuned responses, disruption of which can result in aberrant development or uncontrolled growth, as in cancer. Signaling events are often triggered by posttranslational modifications, such as protein phosphorylation, and subsequent interaction of the modified protein with a binding partner (1). In many cases, signaling proteins are phosphorylated on multiple sites (2). Multiple phosphorylations offer several possible mechanisms to fine-tune signaling events. For example, different kinases may be required for the phosphorylation of different sites on a protein, often by means of “priming” events, in which phosphorylation at one site enables recognition by a subsequent kinase that phosphorylates another site (3). In some instances, multiple phosphorylations can lead to graded regulation of a binding interaction (4). Of particular interest to the present study is the requirement for multiple phosphorylations of a disordered protein by the same kinase to bind to its target. This behavior can lead to a highly sensitive response by transforming a graded input (the kinase concentration, [kinase]) into a sharply thresholded output (5, 6), even in a simple steady-state setting without the presence of ultrasensitive changes in the phosphorylation level entailed by open-system chemistry (7–9). In the simple case with independent phosphorylation sites in the presence of a counteracting phosphatase, this is because the output of an effect that requires  $n$  phosphorylations varies approximately as  $[\text{kinase}]^n$  before the output level is saturated. Hence, when the value of  $n$  is sufficiently larger than one, output stays at an insignificantly low level and increases sharply only when [kinase] attains a certain threshold.

However, the biophysical origin of phosphorylation thresholds, mandating a minimum number of phosphorylations to effect cer-

tain biological events, remains to be deciphered. A mathematical formalism for understanding such processes has been proposed (10), but the underlying physics of molecular interactions has not been elucidated. Here we address the physicochemical forces that might enable a multiphosphorylated protein to bind with high affinity to a partner above a certain phosphorylation threshold, even though the binding affinity of each individual phosphorylated epitope might be very weak.

A model ultrasensitive system that involves a single receptor site and a polyvalent ligand is the interaction of the cyclin-dependent kinase inhibitor Sic1 with the SCF ubiquitin ligase subunit Cdc4 in yeast. When Sic1 is phosphorylated on any six (or more) of nine sites [termed Cdc4 phosphodegrons (CPDs)], it binds to a WD40 domain in Cdc4 with concomitant ubiquitination and degradation of Sic1 (11, 12). The phosphorylation-dependent degradation of Sic1 in the late G<sub>1</sub> phase of the cell cycle enables the development of B-type cyclin–cyclin-dependent kinase activity and the onset of DNA replication (13–15). It has been demonstrated that the phosphodegrons of Sic1 consist of suboptimal binding motifs for binding to the WD40 domain. Although a single high-affinity CPD motif is sufficient for recruitment of Sic1 to Cdc4, because the threshold is absent, premature onset of degradation of Sic1 leads to genome instability (11).

Sic1 is an intrinsically disordered protein (11). Multisite phosphorylation often regulates the function of intrinsically disordered proteins, such as Sic1 (see, e.g., ref. 16); moreover, the sequence patterns of regions adjacent to phosphorylation sites are often very similar to those of intrinsically disordered regions (17). As the key roles of intrinsically disordered proteins in cellular processes are recognized, the study of this class of proteins is flourishing (18–22). Thus, it has become increasingly clear that biology has made use of the manifold polymeric states available to a polypeptide chain (18). Even for proteins that can be folded, in view of recent advances in the study of possible noncooperative “downhill” folding (23–25), the traditional notion of cooperative folding as a functional requirement for globular proteins seems an oversimplification (26). For many tasks involved in cellular function, intrinsically disordered proteins offer several unique attributes. For instance, disordered regions can be flexible linkers or thermodynamic tethers (27) and “entropic springs” (28). Conformational plasticity may also allow such regions to bind to multiple partners and thus serve as hubs in protein–protein interaction networks (29, 30). It has also been suggested that because intrinsically disordered proteins can encode

Author contributions: M.B., T.M., J.D.F.-K., and H.S.C. designed research; M.B., T.M., and H.S.C. performed research; M.T. contributed new reagents/analytic tools; M.B., T.M., T.P., M.T., J.D.F.-K., and H.S.C. analyzed data; and M.B., T.M., T.P., M.T., J.D.F.-K., and H.S.C. wrote the paper.

The authors declare no conflict of interest.

Abbreviation: CPD, Cdc4 phosphodegron.

<sup>¶</sup>To whom correspondence may be addressed. E-mail: chan@arrhenius.med.toronto.edu or pawson@mshri.on.ca.

This article contains supporting information online at [www.pnas.org/cgi/content/full/0702580104/DC1](http://www.pnas.org/cgi/content/full/0702580104/DC1).

© 2007 by The National Academy of Sciences of the USA

larger intermolecular interfaces with smaller protein sizes compared with folded proteins, they may allow for more economical genome and cell sizes (21). By sampling a larger volume, conformational disorder may lead to enhanced capture rates for a protein (31) and serve to better exclude other competing proteins from interacting with its binding partners, as illustrated for example by the entropic barrier formed by the unstructured FG repeats of the nuclear pore (32). Some intrinsically disordered proteins undergo disorder-to-order transitions to adopt stable structures upon binding to their partners (22, 33, 34), but some remain largely disordered, even when bound, or only undergo a limited disorder-to-order transition of a short binding motif (21, 22). Our knowledge of the thermodynamics of disordered proteins is limited. But such knowledge is important to account for these binding mechanisms and their biological implications.

In this study, we investigate electrostatic interactions as a component of the physical mechanism that leads to phosphorylation-dependent ultrasensitive binding of intrinsically disordered ligands. In general, there are more charged residues and fewer hydrophobic residues in intrinsically disordered proteins than in globular proteins (35). Thus, electrostatic interactions are expected to play a more prominent role in conformationally flexible (36), intrinsically disordered proteins than in globular proteins. To emphasize that our proposed electrostatic effects encompass all of the charged phosphorylated sites of the ligand, regardless of whether they are in close proximity to the receptor site, we term these interactions “polyelectrostatic.” We construct a simple analytical mean-field model of the binding of a folded, charged receptor protein to a disordered ligand protein with variable phosphorylation levels. This model rationalizes the behavior of phosphorylation-dependent binding under different solvent and phosphorylation conditions and provides insights into an energetic design of this class of highly cooperative protein–protein interactions. Experimentally determined affinities of Cdc4 for a series of peptide fragments from Sic1 with different net charges are consistent with the predictions of the polyelectrostatic model.

## Results and Discussion

**Statistical Mechanics of Binding.** We first outline a basic formulation for describing the interactions between folded proteins and disordered polyvalent ligands. The present work takes an equilibrium modeling approach to address data from experiments performed under equilibrium conditions, although we note that ultimately nonequilibrium analyses (7–9) will be needed to address processes in living cells. Here we consider a system of  $N_r$  identical receptor molecules and  $N_l$  identical ligands in an aqueous solution of volume  $V$ . To address Cdc4–Sic1 interactions, we focus on the case in which each folded protein molecule (corresponding to Cdc4) contains one binding (receptor) site, whereas each ligand contains  $n$  possible binding regions (corresponding to the CPDs of Sic1) that can bind to the single receptor site on the folded protein. We further assume that, for  $n \geq 1$ , each ligand can bind to at most one receptor protein at a time. This condition holds readily if the concentration of receptor protein is sufficiently low. We note that, although Cdc4 is known to dimerize, dimerization does not affect its affinity for Sic1 because the two binding sites in the dimer are well separated (J. Tang, S. Orlicky, Z. Lin, A. Willems, D. Neculai, D. Ceccarelli, F. Mercurio, B. Shilton, F. Sicheri, and M.T., unpublished data). In other words, we consider the binding process  $r + l \rightleftharpoons r \cdot l$  with the dissociation constant

$$K_d = \frac{[r][l]}{[r \cdot l]}, \quad [1]$$

where  $r$ ,  $l$ , and  $r \cdot l$  represent, respectively, free (unbound) receptor, free ligand, and bound receptor–ligand complex; and  $[\dots]$  denotes concentration.

To relate the dissociation constant to microscopic interactions, we start with the standard statistical mechanical description of this system in terms of a total partition function,  $Q_{\text{tot}}$ , which is a product of contributions from the partition functions for a free disordered ligand protein ( $Q_l^{(f)}$ ), a free folded receptor protein ( $Q_r^{(f)}$ ), and a bound receptor–ligand complex ( $Q_c^{(b)}$ ):

$$Q_{\text{tot}} = \frac{(Q_c^{(b)} V)^{N_c^{(b)}}}{N_c^{(b)}!} \frac{(Q_r^{(f)} V)^{N_r^{(f)}}}{N_r^{(f)}!} \frac{(Q_l^{(f)} V)^{N_l^{(f)}}}{N_l^{(f)}!}, \quad [2]$$

where  $V$  is the volume of the solution in a certain unit (the size of which is immaterial to the present analysis);  $Q_l^{(f)}$ ,  $Q_r^{(f)}$ , and  $Q_c^{(b)}$  are partition functions for the respective individual molecules in the free state or for the complex with a given center-of-mass position;  $N_c^{(b)}$  is the number of bound receptor–ligand complexes;  $N_r^{(f)}$  and  $N_l^{(f)}$  are, respectively, the number of free receptor and free ligand molecules, hence  $N_r^{(f)} = N_r - N_c^{(b)}$ ,  $N_l^{(f)} = N_l - N_c^{(b)}$ ,  $[r] = N_r^{(f)}/V$ ,  $[l] = N_l^{(f)}/V$ , and  $[r \cdot l] = N_c^{(b)}/V$ . Because the conformational diversity of each individual molecule or complex has already been factored into their respective partition functions, the factorials in Eq. 2 account for the fact that the  $N_c^{(b)}$  bound receptor–ligand complexes are indistinguishable and that the same applies to the  $N_r^{(f)}$  free receptors and to the  $N_l^{(f)}$  free ligands.

The total partition function is the sum of statistical weights of all possible configurations of the receptor/ligand system. Accordingly, the free energy of the system  $A = -k_B T \ln Q_{\text{tot}}$ , where  $k_B T$  is Boltzmann’s constant times absolute temperature. Using Eq. 2 for  $Q_{\text{tot}}$ , we determined the most probable physical configuration by minimizing the free energy  $A$  [see [supporting information \(SI\) Appendix 1](#)]. This calculation allows the fraction of ligand bound,  $\theta = N_c^{(b)}/N_l$ , to be expressed in terms of (i) the partition functions  $Q_c^{(b)}$ ,  $Q_r^{(f)}$  and  $Q_l^{(f)}$  for individual bound complex and free molecules, (ii) the receptor to ligand concentration ratio  $\gamma = N_r/N_l$ , and (iii) the total ligand concentration  $\rho_l = N_l/V$ :

$$\theta = \frac{1}{2} \left( 1 + \gamma + \frac{Q_r^{(f)} Q_l^{(f)}}{Q_c^{(b)} \rho_l} \right) - \sqrt{\frac{\left( 1 + \gamma + \frac{Q_r^{(f)} Q_l^{(f)}}{Q_c^{(b)} \rho_l} \right)^2}{4}} - \gamma. \quad [3]$$

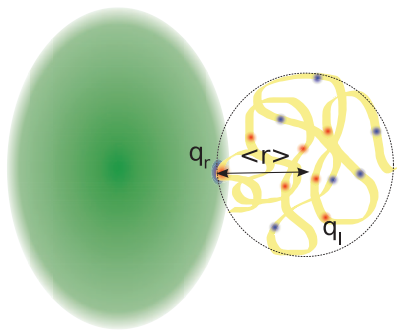
This formulation also provides the concentration of free ligand,  $\rho_l^{(f)} = (1 - \theta)\rho_l$ , and free receptor,  $\rho_r^{(f)} = (\gamma - \theta)\rho_l$ , as functions of the fraction of ligand bound:

$$\rho_l^{(f)} = \frac{Q_r^{(f)} Q_l^{(f)}}{Q_c^{(b)}} \frac{\theta}{(\gamma - \theta)}, \quad \rho_r^{(f)} = \frac{Q_r^{(f)} Q_l^{(f)}}{Q_c^{(b)}} \frac{\theta}{(1 - \theta)}. \quad [4]$$

Because each ligand can bind to only one receptor, the fraction of ligand bound,  $\theta$ , varies between 0 and a maximum value  $\theta_{\text{max}} = \gamma$  for  $\gamma \leq 1$ , and between 0 and  $\theta_{\text{max}} = 1$  for  $\gamma \geq 1$ . Thus, we may normalize  $\theta$  by  $\theta_{\text{max}}$  and consider  $\theta/\theta_{\text{max}}$  as an order parameter for binding. We have computed and compared results for  $\theta_{\text{max}} = 1$  and  $\theta_{\text{max}} < 1$  for several test cases. The general trend of dependence of  $\theta/\theta_{\text{max}}$  on the other variables of the model are very similar (data not shown). For simplicity, numerical results below will be presented for  $\gamma = 1$ . The dissociation constant,  $K_d$ , can now be obtained as the concentration of free receptor,  $\rho_r^{(f)}$ , when  $\theta = 1/2$ :

$$K_d = \frac{Q_r^{(f)} Q_l^{(f)}}{Q_c^{(b)}}, \quad [5]$$

which, when substituted into Eq. 1 with  $[r] = \rho_r(\gamma - \theta)$ ,  $[l] = \rho_l(1 - \theta)$ , and  $[r \cdot l] = \rho_l \theta$ , can readily be checked to yield Eq. 3 for  $\theta$ . Eq. 5 is a general relation valid for all values of  $Q$  and  $\gamma$ . The number of binding regions or phosphorylated sites,  $n$ , of the disordered peptide does not appear explicitly in this equation. We now explore how possible physical interactions determine the values for  $Q$  and consequently lead to a dependence of  $K_d$  on  $n$ .



**Fig. 1.** Schematic of a conformationally disordered ligand (yellow) with both positive (blue) and negative (red) charges (for a total charge,  $q_l$ ) that is bound to a receptor (green) with a binding site having charge  $q_r$ . The distance ( $r$ ) is between the binding site of the receptor and the center of mass of the disordered ligand.

**Model of Electrostatic Interactions.** We examine how the electrostatic interactions between the receptor and the multiple phosphorylated sites might lead to ultrasensitivity. Because the effect of phosphorylation is to add two negative charges to a residue and because electrostatic interactions likely play a significant role (37, 38) in the structural consequences of phosphorylation (39), an obvious candidate for the physical basis of ultrasensitivity is changes in electrostatic interactions. As a first step in our investigation, we consider a physical picture in which the charges of the disordered ligand are dynamically distributed in a diffuse manner, as illustrated in Fig. 1. In other words, we adopt an equilibrium, mean-field perspective (40, 41) in which the electrostatic interactions are not specific, recognizing that such interactions can be biologically functional, for example, in the association of peripheral proteins to cell membranes (42).

Our approach is based on the following simplifying assumptions. (i) The conformational shape and the overall charge distribution of the disordered ligand are not significantly affected by binding. (ii) The electrostatic interaction between the receptor and a bound, disordered ligand can be described approximately as a screened Coulombic interaction between two charges ( $q_r$  on the receptor and the total net charge on the ligand,  $q_l$ ) separated by an effective distance ( $r$ ) (Fig. 1). The present treatment does not distinguish between enthalpic contributions and possible entropic contributions from counterion release (43). (iii) The binding of the disordered ligand and the receptor entails, in addition to the electrostatic interaction, a favorable contact energy  $E_b$  comprised of van der Waals and other energetic terms that derive from the short-range contacts within the binding interface. We further assume that  $E_b$  is independent of whichever one of the  $n$  possible binding regions of the disordered ligand is in contact with the binding site on the receptor. (iv) This favorable contact interaction and the electrostatic interaction are absent when the disordered ligand is unbound.

It follows from assumption *i* that the partition function for a complex is proportional to the product of the partition functions of its two constituent parts,  $Q_c^{(b)} \propto (\delta V) Q_r^{(f)} Q_l^{(f)}$ , because the respective internal degrees of freedom of the receptor and the disordered ligand in the complex are taken to be roughly the same as when they are individually free. The factor ( $\delta V$ ) is a volume of tolerance delineating the range of variation of the separation between the centers of mass of the two constituent molecules in the complex (44). This factor sets the volume (and concentration) scale in the formulation. Assumptions *ii* and *iii* further imply that

$$Q_c^{(b)} \propto n \exp\left(-\frac{E_b + q_r q_l e^{-\alpha(r)/\epsilon_d \langle r \rangle}}{k_B T}\right) \quad [6]$$

where  $\epsilon_d$  is the dielectric constant and  $\alpha$  is a Debye–Hückel screening parameter. The number  $n$  of possible binding regions on

the ligand appears as an overall multiplicative factor because it corresponds to the number of distinguishable ways of forming the complex, assuming that, to a first approximation, they are equally likely to bind. Thus, by combining Eqs. 5 and 6, it follows that the dissociation constant

$$K_d = \frac{Q_r^{(f)} Q_l^{(f)}}{Q_c^{(b)}} = \frac{1}{(n \delta V)} \exp\left(\frac{E_b + q_r q_l e^{-\alpha(r)/\epsilon_d \langle r \rangle}}{k_B T}\right), \quad [7]$$

which depends on  $n$  explicitly in the overall  $1/(n \delta V)$  factor as well as implicitly by means of  $q_l$  in the exponential factor (see below).

Obviously, this formulation is a highly simplified caricature of the real system. For instance, in reality, the conformational shape of the ligand could change upon binding, and the electrostatic interactions are more accurately described by the Coulombic attractions and repulsion between individual charges than by net charges and an effective separation  $\langle r \rangle$ . Incorporation of these effects would likely lead to tighter binding, as discussed below and in *SI Appendix 1*. The present treatment also assumes that receptor interaction with different binding sites of the disordered ligand entail the same favorable contact energy  $E_b$ . However, in reality it is known that the CPD sequences on Sic1 differ from one another, with various degrees of mismatch to the canonical CPD consensus sequence (11). In addition, the current model addresses only the thermodynamics but not the kinetics of binding (10). Assuming no interaction between the receptor and the ligand in the unbound state means that there is little or no kinetic driving forces for the binding interaction other than diffusion; this might not be realistic. Despite these limitations, in our estimation, this model captures much of the essential physics of interest. It also offers a practical means to study qualitatively or even semiquantitatively the interplay between ligand net charge  $q_l$  (see below) and other solvent variables. For example, the effects of salt concentration on binding may be explored by noting that  $\alpha \propto \sqrt{I}$  where  $I$  is the ionic strength of the solution, as in previous mean-field theories of biopolymers (40). In this regard, to a first approximation, addition of salt is expected to weaken the polyelectrostatic interactions.

**Ultrasensitivity as a Consequence of Cumulative Ligand–Receptor Electrostatic Interactions.** Based on this modeling setup, Fig. 2 examines the origin of ultrasensitivity in our model. The number of phosphorylation sites enters into our formalism (*i*) as an overall multiplicative factor for the partition function  $Q_c^{(b)}$  of the bound complex, as noted above, and (*ii*) as a multiplicative factor on a negative contribution to the net charge  $q_l$  of the ligand in Eq. 6:

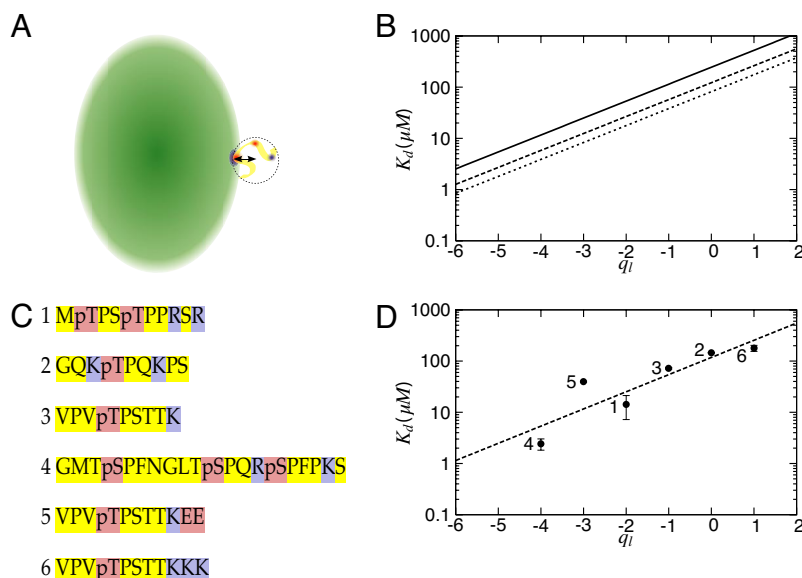
$$q_l = q_l^0 - c n, \quad [8]$$

where  $q_l^0$  is the charge of the unphosphorylated ligand and  $c$  is the charge contribution per phosphorylation;  $c$  is taken to be 2 at pH 7. The strength of the electrostatic interaction depends on the dielectric constant  $\epsilon_d$ , which is  $\approx 80$  for bulk water and approximately 2–4 for the interior of a folded protein molecule (45). Because the ligand is disordered in both the free and bound state, the numerical results in Fig. 2 are presented for an intermediate value of  $\epsilon_d = 20$ . We have performed calculations by using other values of  $\epsilon_d$  ( $5 \leq \epsilon_d \leq 80$ ); the general trend of the results are similar (data not shown).

Fig. 2*A* depicts ultrasensitive behavior in our model. The fraction  $\theta$  of ligands bound undergoes a steep sigmoidal transition from  $\theta \approx 0$  at  $n = 5$  to  $\theta \approx 1$  at  $n = 7$  or 8. Hence, it is a sharp binding transition that is triggered by a threshold value of  $n$  and also displays switch-like properties (46). The threshold  $n$  value increases with decreasing total ligand concentration,  $\rho_l$ . However, the dependence of threshold  $n$  value on  $\rho_l$  is not strong. In this example, the threshold  $n$  changes only from 6 to 7 for a variation of  $\rho_l$  over two orders of magnitude. Although the exact threshold  $n$  value in the model depends on the choice of parameters, Fig. 2*A* demonstrates







**Fig. 4.** Sic1 peptide binding to Cdc4. (A) Schematic of a short peptide fragment bound to a receptor protein. (B) Binding constants vs. ligand net charge ( $q_l$ ) in the mean-field model, with  $q_r = 3.0$ ,  $E_b = -15 k_B T$ ,  $\langle r \rangle = 12 \text{ \AA}$ , and  $\epsilon_d = 44$ .  $K_d$  values are shown for  $n = 1$  (solid line), 2 (dashed line), and 3 (dotted line). (C) Peptide fragment sequences. Blue (red) background corresponds to positively (negatively) charged residues. pS and pT denote phosphorylated serine and threonine, respectively. (D) Experimental dissociation constants vs. ligand net charge for peptide fragments shown in C. The line is a least-square fit of  $\log(K_d)$  vs.  $q_l$ . Error bars are one standard deviation of the results from three independent titrations. The number labels for the fragment sequences correspond to those in C.

Accordingly, theoretical binding curves for  $n = 1, 2$ , or 3 are provided in Fig. 4B, which shows that, for a given  $q_l$ , the variation in  $K_d$  among the plotted  $n$  values is quite small. The experimental binding affinities as a function of peptide net charge are shown in Fig. 4D. The dependence of  $K_d$  on the net charge is to a good approximation exponential (linear in the logarithmic scale shown), which is in close agreement with the model predictions in Fig. 4B.

These experimental results lend support to our hypothesis that polyelectrostatic interactions are a significant driving force for ultrasensitivity in the Sic1/Cdc4 system. In this regard, it should be noted that although the model  $\delta V$  parameter in Fig. 4B was tuned to mimic the experimental results, the linear dependence of  $\log K_d$  vs. net charge is a robust feature of the model. From the slope of  $\log K_d$  vs. Sic1 peptide net charge in Fig. 4D, we may estimate a binding free energy change of  $0.77 k_B T$  per unit charge. This estimation means a free energy change of  $1.54 k_B T$ , and hence an  $\approx 4.7$ -fold increase in  $K_d$  per phosphorylation. Assuming a similar charge dependence for the binding of 90-residue Sic1 (1–90), this significant difference in predicted affinity between  $n$ -fold and  $(n + 1)$ -fold phosphorylated peptide is consistent with the experimental findings of significant degradation of 6-fold but not 5-fold phosphorylated Sic1 (12) and the ability of Cdc4 to “capture” 6-fold but not 5-fold phosphorylated Sic1 (11). The relationship between  $K_d$  and the ligand net charge also depends on the detailed charge distribution and ligand length. In *SI Appendix 1*, we explore this relationship in terms of polyampholyte polarizability and show how longer ligands can compensate for unfavorable electrostatic interaction by conformational rearrangements and that this effect depends on charge distribution and ligand flexibility.

## Conclusions

We have investigated the mathematical and physical requirements for ultrasensitivity in a simple polyvalent ligand–single receptor system. The observation of a switch-like change in binding affinity places constraints on how the population of bound complex or the fraction of ligand bound,  $\theta$ , depends on the number,  $n$ , of phosphorylated sites. It is not sufficient for  $\theta$  to increase linearly with  $n$ . But ultrasensitivity can be a consequence of  $\theta$  increasing exponentially with  $n$  (Fig. 2). This requirement is formally similar to that

for other cooperative transitions in biology, such as protein folding (26).

Such an exponential dependence was also postulated in a previous mathematical model for ultrasensitivity (10), although this earlier study did not provide a physical mechanism. A recent proposal argues that multiple phosphorylation events can in principle favor binding by reducing the conformational entropy of a disordered ligand when binding is coupled with a high degree of conformational ordering of the ligand (49). However, for the Sic1–Cdc4 system, the available experimental evidence suggests that phosphorylation on multiple sites does not lead to substantial decrease in the conformational disorder of Sic1 (11) (T.M., W.-Y. Choy, S. Orlicky, F. Sicheri, L. Kay, M.T., and J.D.F.-K., unpublished data).

Noting the obvious fact that phosphorylation increases negative charges on a ligand, here we show that a viable physical mechanism for ultrasensitivity is readily supplied by the increasingly favorable electrostatic interactions between a positively charged receptor protein and a disordered ligand protein with an increasing number of phosphorylated sites. This formulation satisfies the above mathematical condition because the electrostatic interaction energy naturally enters into the exponential factor of the Boltzmann weights. Our hypothesis appears consistent with existing data and the experimental data reported in this work.

The present model highlights the important interplay between the more specific, spatially short-range contact interactions in and around the binding pocket and the less specific, spatially long-range polyelectrostatic interactions between the binding partners. Consistent with experiments on the Sic1–Cdc4 system (11), polyelectrostatic interactions can lead to ultrasensitivity only when the strength of the specific contact interactions ( $E_b$  in our model) is not too strong (Fig. 3). In other words, the specific contact interactions must be insufficient for binding in the first place such that cumulative charge interactions, e.g., by increasing the number of phosphorylations for the Sic1/Cdc4 case, can tip the balance toward favorable binding. Because our model is rather generic, this effect should be applicable to a broad class of polyvalent intrinsically disordered ligands. For example, the F-box protein Grr1 engages in a multisite phosphorylation-dependent interaction with a disor-

dered region in its substrate the G<sub>1</sub> cyclin Cln2 (50) and a phosphorylation-independent interaction with an acidic degron sequence of another substrate, Mks1 (51). Notably, and unlike Cdc4, Grr1 does not appear to have a dedicated phosphopeptide binding pocket but instead interacts with substrates by means of a cationic surface on its leucine-rich repeat domain (52). These disordered polyelectrostatic interactions contrast with interactions involving a disorder-to-order transition upon binding (22), for which the specific contact interactions involving the phosphate leads to significant structural stabilization. In addition, there are certain to be other mechanisms of binding interactions involving phosphorylation of disordered proteins that do not lead to ultrasensitivity, as in the phosphorylation of an unstructured region of the transcription activator Ets-1, which results in a graded DNA binding affinity essentially as a “rheostat” rather than a switch (4).

Finally, we emphasize that the polyelectrostatic perspective put forth in this work is more general than the simple mean-field model used to illustrate it here. The present treatment considers only the screened electrostatic interaction between a pair of effective point charges, one representing the receptor, the other representing the ligand, without accounting for charge distribution. However, it is also clear from the formal development that the trends predicted by our model are readily generalizable, e.g., to cases for which there are multiple specific contacts instead of only one specific contact region between the intrinsically disordered ligand protein and the receptor protein. By incorporating more structural and energetic details, our perspective can potentially address many more subtle phenomena, as exemplified by the brief exploration of the impact of polyampholytic charge polarization (53–55) in the disordered ligand (*SI Appendix 1*). In this context, it would be fruitful to investigate the effects of charge distribution on the receptor side as well. There is no shortage of intriguing fundamental biological

questions for which the polyelectrostatic perspective would be of relevance.

## Experimental Procedures

The binding measurements of Sic1 fragments to Cdc4 were carried out by using an ATF 105 spectrometer (AVIV Instruments, Lakewood, NJ) equipped with an automatic titrator. The binding of the peptide fragments were monitored by using intrinsic tryptophan fluorescence with an excitation wavelength of 298 nm and an emission wavelength of 366 nm. All measurements were performed in 50 mM Hepes/150 mM NaCl/5 mM DTT, pH 7.5, at 22°C. The concentration of Cdc4 was 0.5  $\mu$ M for all measurements. The Cdc4 concentrations were determined by using UV absorption at 280 nm in 6 M guanidine hydrochloride. The concentrations of the peptide fragments were assessed by BCA assays and amino acid analysis.

Cdc4 used for binding assays was the same monomeric Cdc4–Skp1 complex used for crystallography with Cdc4 encompassing residues 263–744 (12). It was expressed and purified as reported before (12). Phosphopeptides were obtained from Genscript and the Advanced Protein Technology Centre at the Hospital for Sick Children, and purified by reversed-phase chromatography on a C18 column. The identity and proper phosphorylation of the peptides were confirmed by amino acid analysis and MALDI mass spectrometry.

We thank James Wing-Yiu Choy, Hong Qian, and Elliott Stollar for helpful discussions and Alan Davidson for use of his fluorimeter. This work was supported in part by grants from the National Cancer Institute of Canada (to M.T. and J.D.F.-K.) and Canadian Institutes of Health Research (to J.D.F.-K. and H.S.C.). M.B. was supported in part by the Canadian Institutes of Health Research Training Program in Protein Folding at the University of Toronto and a postdoctoral fellowship from the Sweden–America Foundation. T.M. is the recipient of a Terry Fox postdoctoral research fellowship. M.T. and H.S.C. are Canada Research Chair holders.

- Pawson T, Nash P (2003) *Science* 300:445–452.
- Cohen P (2000) *Trends Biochem Sci* 25:596–601.
- Frame S, Cohen P (2001) *Biochem J* 359:1–16.
- Pufall MA, Lee GM, Nelson ML, Kang HS, Velyvis A, Kay LE, McIntosh LP, Graves BJ (2005) *Science* 309:142–145.
- Ferrell JE, Jr (1996) *Trends Biochem Sci* 21:460–466.
- Salazar C, Höfer T (2003) *J Mol Biol* 327:31–45.
- Goldbeter A, Koshland DE, Jr (1981) *Proc Natl Acad Sci USA* 78:6840–6844.
- Angeli D, Ferrell JE, Jr, Sontag ED (2004) *Proc Natl Acad Sci USA* 101:1822–1827.
- Qian H (2007) *Annu Rev Phys Chem* 58:113–142.
- Klein P, Pawson T, Tyers M (2003) *Curr Biol* 13:1669–1678.
- Nash P, Tang X, Orlicky S, Chen Q, Gertler FB, Mendenhall MD, Sicheri F, Pawson T, Tyers M (2001) *Nature* 414:514–521.
- Orlicky S, Tang X, Willems A, Tyers M, Sicheri F (2003) *Cell* 112:243–256.
- Schwob E, Böhm T, Mendenhall MD, Nasmyth K (1994) *Cell* 79:233–244.
- Tyers M (1996) *Proc Natl Acad Sci USA* 93:7772–7776.
- Verma R, Annan RS, Huddleston MJ, Carr SA, Reynard G, Deshaies RJ (1997) *Science* 278:455–460.
- Honnappa S, Jahnke W, Seelig J, Steinmetz MO (2006) *J Biol Chem* 281:16078–16083.
- Lakoucheva LM, Radivojac P, Brown CJ, O'Connor TR, Sikes JG, Obradovic Z, Dunker AK (2004) *Nucl Acids Res* 32:1037–1049.
- Dunker AK, Lawson JD, Brown CJ, Williams RM, Romero P, Oh JS, Oldfield CJ, Campen AM, Ratliff CR, Hipps KW, et al. (2001) *J Mol Graphics Model* 19:26–59.
- Tomba P (2002) *Trends Biochem Sci* 27:527–533.
- Uversky VN (2002) *Eur J Biochem* 269:2–12.
- Gunasekaran K, Tsai CJ, Kumar S, Zanuy D, Nussinov R (2003) *Trends Biochem Sci* 28:81–85.
- Dyson HJ, Wright PE (2005) *Nat Rev Mol Cell Biol* 6:197–208.
- Sadqi M, Fushman D, Munoz V (2006) *Nature* 442:317–321.
- Knott M, Chan HS (2006) *Proteins* 65:373–391.
- Huang F, Sato S, Sharpe TD, Ying L, Fersht AR (2007) *Proc Natl Acad Sci USA* 104:123–127.
- Chan HS, Shimizu S, Kaya H (2004) *Methods Enzymol* 380:350–379.
- Bowman P, Galea CA, Lacy E, Kriwacki RW (2006) *Biochim Biophys Acta* 1764:182–189.
- Rauscher S, Baud S, Miao M, Keeley FW, Pomes R (2006) *Structure (London)* 14:1667–1676.
- Dunker AK, Cortese MS, Romero P, Lakoucheva LM, Uversky VN (2005) *FEBS J* 272:5129–5148.
- Patil A, Nakamura H (2006) *FEBS Lett* 580:2041–2045.
- Shoemaker BA, Portman JJ, Wolynes PG (2000) *Proc Natl Acad Sci USA* 97:8868–8873.
- Lim RYH, Huang NP, Koser J, Deng J, Lau KHA, Schwarz-Herion K, Fahrenkrog B, Aebi U (2006) *Proc Natl Acad Sci USA* 103:9512–9517.
- Verkhivker GM (2005) *Proteins* 58:706–716.
- Wang J, Lu Q, Lu HP (2006) *PLoS Comput Biol* 2:e78.
- Romero P, Obradovic Z, Li X, Garner EC, Brown CJ, Dunker AK (2001) *Proteins* 42:38–48.
- Radivojac P, Obradovic Z, Smith DK, Zhu G, Vucetic S, Brown CJ, Lawson D, Dunker AK (2004) *Protein Sci* 13:71–80.
- Andrew CD, Warwicker J, Jones GR, Doig AJ (2002) *Biochemistry* 41:1897–1905.
- Zor T, Mayr BM, Dyson HJ, Montminy MR, Wright PE (2002) *J Biol Chem* 277:42241–42248.
- Johnson LN, Barford D (1993) *Annu Rev Biophys Biomol Struct* 22:199–232.
- Stigter D, Dill KA (1990) *Biochemistry* 29:1262–1271.
- Dill KA, Stigter D (1995) *Adv Protein Chem* 46:59–104.
- Murray D, Ben-Tal N, Honig B, McLaughlin S (1997) *Structure (London)* 5:985–989.
- Olmsted MC, Bond JP, Anderson CF, Record MT, Jr (1995) *Biophys J* 68:634–647.
- Qian H (2006) *J Math Biol* 52:277–289.
- Leach AR (1996) *Molecular Modeling: Principles and Applications*. (Longman, Singapore).
- Gunawardena J (2005) *Proc Natl Acad Sci USA* 102:14617–14622.
- Cantor CR, Schimmel PR (1980) *Biophysical Chemistry* (Freeman, New York).
- Carlsson F, Linse P, Malmsten M (2001) *J Phys Chem B* 105:9040–9049.
- Lenz P, Swain PS (2006) *Curr Biol* 16:2150–2155.
- Lanker S, Valdivieso MH, Wittenberg C (1996) *Science* 271:1597–1601.
- Liu X, Spirek M, Thornton J, Butow RA (2005) *Mol Biol Cell* 16:4893–4904.
- Hsiung YG, Chang HC, Pellequer JL, La Valle R, Lanker S, Wittenberg C (2001) *Mol Cell Biol* 21:2506–2520.
- Doty P, Imahori K, Klemperer E (1958) *Proc Natl Acad Sci USA* 44:424–431.
- Higgs PG, Joanny JF (1991) *J Chem Phys* 94:1543–1554.
- Dobrynin AV (2001) *Phys Rev E* 63:051802.